Received: 30 September 2020 | Accepted: 23 February 2021

DOI: 10.1002/ajb2.1710

RESEARCH ARTICLE

American Journal of Botany

Microsatellites and petal morphology reveal new patterns of admixture in *Orchis* hybrid zones

Leif Bersweden[1,2] | Juan Viruel[1] | Bertrand Schatz[3] | Joanna Harland[4] | Roberta Gargiulo[1] | Robyn S. Cowan[1] | Jacopo Calevo[5] | Ana Juan[6] | James J. Clarkson[1] | Andrew R. Leitch[2] | Michael F. Fay[1,7]

[1]Jodrell Laboratory, Royal Botanic Gardens, Kew TW9 3DS, UK

[2]School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

[3]Centre for Ecology and Evolution, University of Montpellier, Montpellier 34090, France

[4]Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

[5]Department of Life Sciences and Systems Biology, University of Turin, Turin 10125, Italy

[6]Department of Environmental Sciences & Natural Resources, University of Alicante, San Vicente, Alicante 03690, Spain

[7]School of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia

Correspondence
Leif Bersweden, Jodrell Laboratory, Royal Botanic Gardens, Kew TW9 3DS, UK.
Email: l.bersweden@kew.org

Michael F. Fay, Jodrell Laboratory, Royal Botanic Gardens, Kew TW9 3DS, UK.
Email: M.Fay@kew.org

Abstract

Premise: The genetic structure of hybrid zones provides insight into the potential for gene flow to occur between plant taxa. Four closely related European orchid species (*Orchis anthropophora*, *O. militaris*, *O. purpurea*, and *O. simia*) hybridize when they co-occur. We aimed to characterize patterns of hybridization in *O. militaris*–*O. purpurea*, *O. purpurea*–*O. simia*, and *O. anthropophora*–*O. simia* hybrid zones using molecular and morphological data.

Methods: We used 11 newly isolated nuclear microsatellites to genotype 695 individuals collected from seven hybrid zones and six allopatric parental populations in France. Geometric morphometric analysis was conducted using 15 labellum landmarks to capture the main aspects of petal shape.

Results: Backcrossing was asymmetric toward *O. militaris* in multiple *O. militaris*–*O. purpurea* hybrid zones. Hybrids in *O. purpurea*–*O. simia* and *O. anthropophora*–*O. simia* hybrid zones were largely limited to F1 and F2 generations, but further admixture had occurred. These patterns were reflected in labellum geometric morphometric data, which correlated strongly with nuclear microsatellite data in all three species combinations.

Conclusions: The coexistence of parental and admixed individuals in these *Orchis* hybrid zones implies they are likely to be tension zones being maintained by a balance between gene flow into the hybrid zone and selection acting against admixed individuals. The pattern of admixture in the three species combinations suggests intrinsic selection acting on the hybrids is weaker in more closely related taxa.

KEYWORDS

backcrossing, genetic structure, geometric morphometrics, hybrid index, hybrid zone, hybridization, nuclear microsatellites, *Orchis*, Orchidaceae, tension zone

Patterns of population genetic structure in threatened species are influenced by a range of factors including climatic fluctuations, habitat fragmentation, and genetic admixture with closely related species (Heuertz et al., 2004). Estimating genetic diversity and structure in species of conservation concern is an important step in implementing informed conservation management, particularly where hybridization and gene flow are occurring. Hybridization between species can lead to introgression, potentially resulting in populations losing their unique genetic identities (Kim et al., 2017).

Studying hybrid zones provides an insight into how species identity is maintained despite ongoing gene flow (Abbott and Brennan, 2014). Three models of hybrid zone structure have been recognized: tension zones (e.g., Brennan et al., 2009), bounded hybrid superiority zones (e.g., Watano et al., 2004), and mosaic hybrid zones (e.g.,

Raudnitschka et al., 2007). Each is defined by the way in which selection acts on individuals to maintain the hybrid zone (Arnold, 1999), which in turn influences the likelihood of gene flow occurring (Abbott and Brennan, 2014).

The maintenance of species boundaries in hybrid zones is dependent on reproductive barriers contributing to reproductive isolation between sexually compatible species growing in sympatry (Coyne and Orr, 2004; Mota et al., 2019). These barriers can be categorized into pre-zygotic (e.g., different pollinator communities, emitted scent, flowering phenology, habitat separation or geographic distribution) and post-zygotic barriers (e.g., hybrid sterility, fruit abortion, mycorrhizal incompatibility or seed inviability), depending on whether they act to avoid interspecific pollination or to prevent the formation of viable offspring (Abbott et al., 2013; Yan et al., 2019).

Hybridization has been widely reported in Orchidaceae, including between food-deceptive orchid species of the genus *Orchis* (Orchidaceae) (Cozzolino et al., 2004; Cozzolino and Widmer, 2005; Kretzschmar et al., 2007). Reproductive isolation is weak (Jacquemyn et al., 2012b), which is unsurprising given the frequency with which hybrid zones form in this group (Kretzschmar et al., 2007). Studies contrasting pre- and post-zygotic reproductive barriers in food-deceptive orchids have shown that pre-zygotic barriers are often relatively insignificant compared to post-zygotic barriers (Schatz, 2006; Scopece et al., 2007, 2008). This is in part because their pollinator communities often consist of generalist pollinators like bees, beetles, and flies that pollinate many different orchid species (Cozzolino and Widmer, 2005; Joffard et al., 2019; Schatz et al., 2020). Experimental crosses have shown that post-zygotic barriers are often strong and prevent further admixture between species (Scopece et al., 2008). Both early-acting (fruit abortion) and late-acting (seed inviability and hybrid sterility) post-zygotic barriers have been shown to play important roles in maintaining species integrity between food-deceptive orchid species (Scopece et al., 2008).

Given that pre-zygotic reproductive barriers appear to be weak or absent in *Orchis*, it can be expected that post-zygotic barriers are strong enough to limit hybridization to early generations, thus restricting introgression and maintaining species integrity. A range of molecular markers have been used to study genetic admixture in *Orchis*, and few studies present evidence of hybridization beyond the first generation (Cozzolino and Widmer, 2005). Analyses of various *Orchis* hybrids, e.g., between *O. anthropophora* and *O. italica* (Pellegrino et al., 2009), *O. purpurea* and *O. simia* (Bateman et al., 2008) and between *O. mascula* and *O. provincialis* (Pellegrino et al., 2005) have shown that most putative hybrids belong to the F1 generation.

Multiple generations of hybrids have been observed in only four species combinations across the entire genus *Orchis* (*O. anatolica*–*O. quadripunctata*, *O. militaris*–*O. purpurea*, *O. militaris*–*O. simia*, and *O. mascula*–*O. pauciflora*) (Cozzolino and Widmer, 2005; Kretzschmar et al., 2007; Jacquemyn et al., 2012a). Studied *Orchis* hybrid zones have never been shown to be entirely composed of hybrid individuals, however, suggesting that post-zygotic barriers are strong enough to maintain species integrity, despite taxa not being completely reproductively isolated.

Data from four closely related diploid *Orchis* species in *Orchis* subgenus *Orchis* (*O. anthropophora*, *O. militaris*, *O. purpurea*, and *O. simia*) provide convincing evidence that hybridization has occurred in the past and is still occurring (Fay et al., 2007). These four species are described as being "anthropomorphic", because they have flowers with a lobed lip that resemble a small human-like figure with distinct "arms" and "legs" (Figures 1 and 2). They have similar geographic distributions, and they can be found growing in sympatry across most of their range, resulting in many local, isolated hybrid zones (Kretzschmar et al., 2007).

To date, only two hybrid zones have been studied in this group of four species: one *O. purpurea*–*O. simia* hybrid zone in the UK (Bateman et al., 2008) and an *O. militaris*–*O. purpurea* hybrid zone in Belgium (Jacquemyn et al., 2012a). We aimed to characterize hybrid zone structure in multiple *O. militaris*–*O. purpurea*, *O. purpurea*–*O. simia* and *O. anthropophora*–*O. simia* hybrid zones in France using nuclear microsatellites and geometric morphometrics.

MATERIALS AND METHODS

Study species

Orchis anthropophora, *O. militaris*, *O. purpurea*, and *O. simia* are tuberous, perennial orchid species belonging to *Orchis* subgenus *Orchis*, otherwise known as the anthropomorphic group of *Orchis* species (Figure 1). *Orchis anthropophora*, *O. purpurea*, and *O. simia* are predominantly associated with the Mediterranean region and western Europe (Kretzschmar et al., 2007). All three species reach the northern boundary of their distributions in the United Kingdom, Belgium, the Netherlands, and Denmark (Willems and Ellers, 1996; Kretzschmar et al., 2007). *Orchis militaris* has a much wider distribution, ranging from western Europe to Mongolia (Farrell, 1985; Kretzschmar et al., 2007). It can also be found further north than the other three species, in Finland, Sweden, and the Baltic states (Kretzschmar et al., 2007; Ilves et al., 2015).

Floral morphology can be used to distinguish these orchid species, particularly the shape and coloration of the labellum (Figures 1 and 2). They have overlapping habitat preferences, flowering phenology, and pollinator communities (Appendix S1).

Plant material used for developing novel microsatellite markers

Genomic libraries were prepared using high-molecular-weight genomic DNA from 11 *Orchis* samples representing both subgenera (Appendix S2: Table S1). To test whether the developed markers worked across different species and

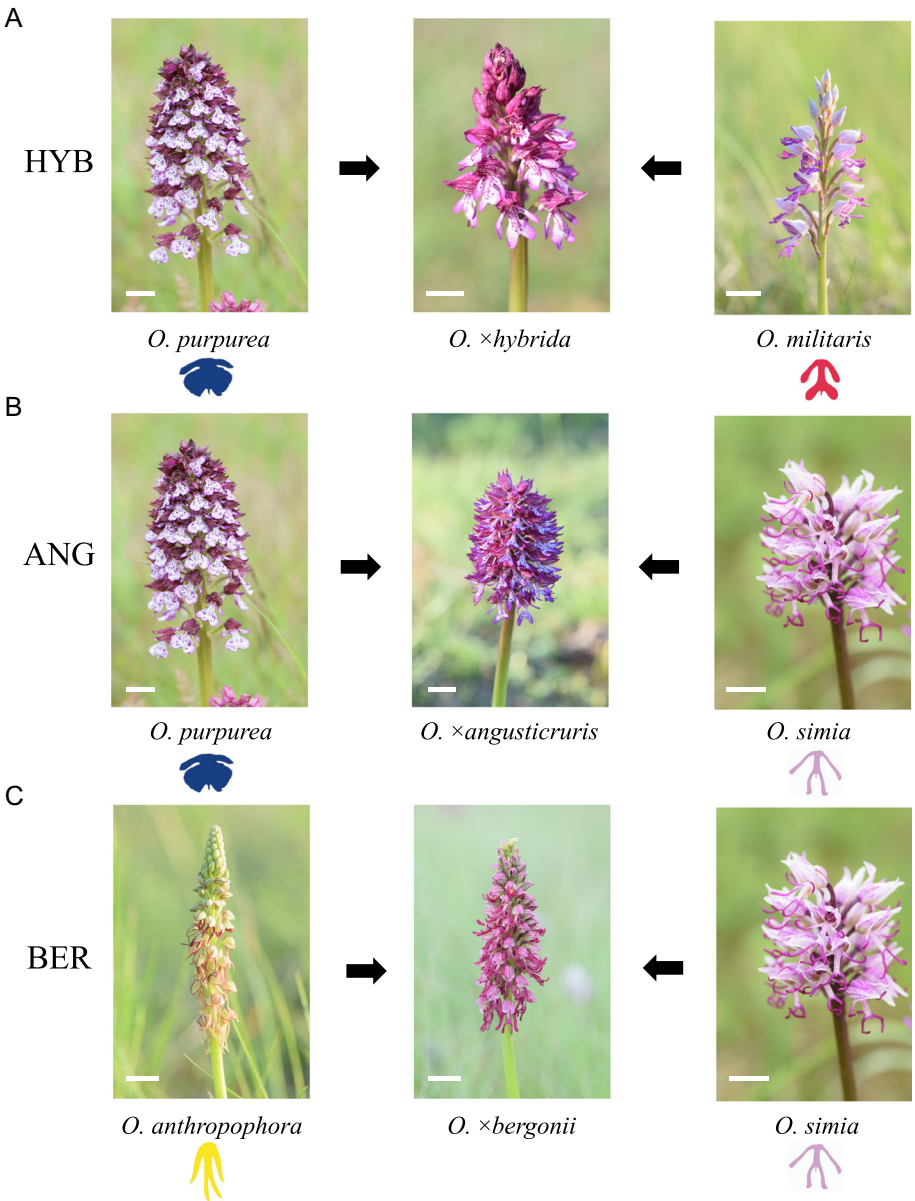

FIGURE 1 (A) HYB populations comprise *O. purpurea*, *O. militaris*, and the hybrid *O.* ×*hybrida*; (B) ANG populations comprise *O. purpurea*, *O. simia*, and the hybrid *O.* ×*angusticruris*; (C) BER populations comprise *O. anthropophora*, *O. simia*, and the hybrid *O.* ×*bergonii*. Internal scale bars are all 10 mm. Representative labella and colors will be used throughout. All photographs by Leif Bersweden

were sufficiently polymorphic within species, we used samples previously extracted and stored in the DNA Bank at the Royal Botanic Gardens, Kew (RBG, Kew; Appendix S2: Table S2). These samples were collected from a wide geographic area to represent the European distribution of the four taxa and collected from populations that were not in sympatry with any other *Orchis* species, so they are not subject to contemporary introgression.

Hybrid populations and DNA extraction

In 2018 and 2019, we collected floral material from seven sympatric populations or hybrid zones in southern France: four *O. militaris*–*O. purpurea* populations with the hybrid *O.* ×*hybrida* (herein referred to as HYB populations; Figure 1A), two *O. purpurea*–*O. simia* populations and the hybrid *O.* ×*angusticruris* (referred to as ANG populations; Figure 1B) and one *O. anthropophora*–*O. simia* population with the hybrid *O.* ×*bergonii* (referred to as the BER population; Figure 1C) (Table 1). Given the significant overlap in geographic distribution between the four species and that hybridization is not a result of large parental populations meeting, each sympatric population can be considered a separate, localized hybrid zone. All seven sites are in calcareous grassland on the limestone plateaus of two protected areas: the Natural Regional Park of Grands Causses and the Cévennes National Park.

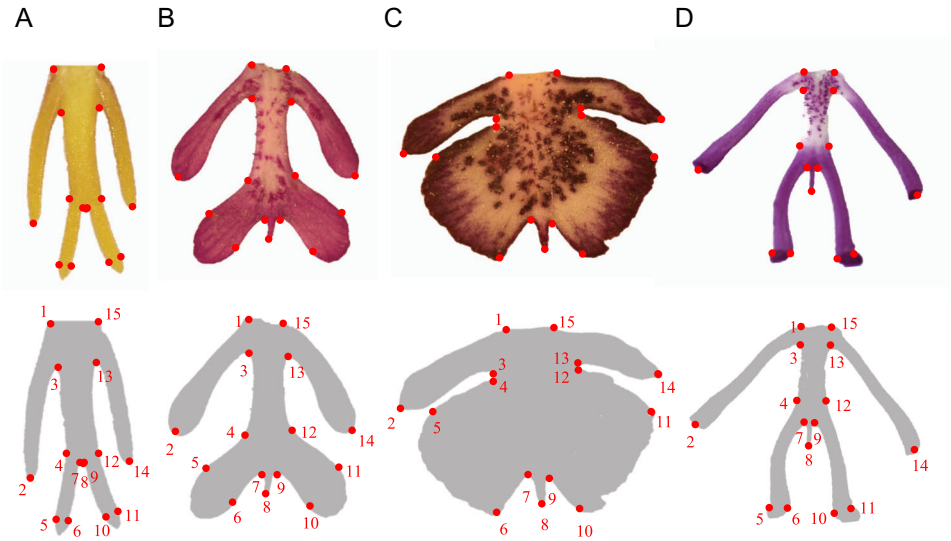

FIGURE 2 Relative positions of the 15 homologous landmarks chosen for geometric morphometric analysis of anthropomorphic *Orchis* species. (A) *O. anthropophora*, (B) *O. militaris*, (C) *O. purpurea*, and (D) *O. simia*. Anthropomorphic terminology used to describe shape changes: arms (landmarks 1–3 and 13–15), torso (landmarks 3–4 and 12–13), legs (landmarks 4–7 and 9–12) and tail (landmarks 7–9)

TABLE 1 *Orchis* floral material collected from southern France for DNA analysis

Population	Species present	Location	*N*
ANG1	*O. purpurea–O. simia*	Alzon (Gard)	30
ANG2	*O. purpurea–O. simia*	Trèves (Gard)	63
BER1	*O. anthropophora–O. simia*	Revens (Aveyron)	36
HYB1	*O. militaris–O. purpurea*	La Pezade (Aveyron)	85
HYB2	*O. militaris–O. purpurea*	Sauclières (Gard)	90
HYB3	*O. militaris–O. purpurea*	Veyreau (Aveyron)	83
HYB4	*O. militaris–O. purpurea*	La Trivalle (Aveyron)	128
OA1	*O. anthropophora*	Lignairolles (Aude)	30
OM1	*O. militaris*	Cantobre (Aveyron)	30
OM2	*O. militaris*	Col de Perjuret (Lozère)	30
OP1	*O. purpurea*	Manses (Ariège)	30
OP2	*O. purpurea*	Les Menudes (Aveyron)	30
OS1	*O. simia*	Blandas (Gard)	30

The number of samples obtained from each population ranged from 30 to 128 (total $N = 695$), and efforts were made in each case to sample equal numbers of parental taxa based on morphology. Morphological factors taken into consideration included flower color, labellum shape, plant height, and inflorescence development (acro- or basipetal). Floral material was also collected from allopatric populations representing the four species ($N = 30$ for each; Table 1). All samples were stored in silica gel (Chase and Hills, 1991). Genomic DNA was extracted using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), and concentrations were measured using a Quantus fluorometer (Promega, Madison, WI, USA).

DNA library preparation and genome skim sequencing

Fragment size was estimated using gel electrophoresis, and samples with high-molecular-weight DNA were chosen (Appendix S2: Table S1). Samples were diluted to 50 ng μL^{-1} using UltraPure 10 mM Tris-HCl, pH 8.0 and sonicated with a Covaris ME220 Focused-ultrasonicator (Covaris, Woburn, MA, USA) to generate fragments ca. 300 base pairs (bp) long.

Genomic libraries were prepared from the fragmented genomic DNA according to the NEBNext Ultra II DNA Library Prep Kit for Illumina protocol (New England BioLabs, Ipswich, MA, USA). AMPure XP beads (Beckman Coulter Life Sciences, Indianapolis, IN, USA) were used for size selection (300–350 bp). Each genomic library was subjected to eight cycles of PCR amplification, their fragment size was measured using an Agilent 4200 TapeStation (Agilent Technologies, Santa Clara, CA, USA), and DNA concentration was quantified using a Quantus fluorometer (Promega, Madison, WI, USA). Sequencing was undertaken at RBG, Kew on an Illumina MiSeq (Illumina Inc., San Diego, CA, USA). Sequence reads were filtered for read quality using FASTQC v. 0.11.5 (Andrews, 2010) and trimmed using Trimmomatic v. 0.39 (Bolger et al., 2014).

Microsatellite marker screening

On average, 1,907,570 reads were obtained per sample. MSATCOMMANDER v. 1.0.8 (Faircloth, 2008) was used to

identify microsatellites using the following options: 50–350 bp PCR products, with a minimum of five repeats of 2–4-bp motifs and 18–23-bp primer length. The number of novel microsatellite loci retrieved per sample using MSATCOMMANDER ranged from 5,399 to 16,480, with a total of 99,116 (Rozen and Skaletsky, 2000; Faircloth, 2008). The following factors were taken into consideration when manually screening the microsatellites for suitable loci (Viruel et al., 2018): (1) low self- or pair-complementarity, (2) <1.0°C difference in melting temperature between forward and reverse primers, (3) high number of repeats per locus to increase the likelihood of there being variation, (4) high GC-content to increase primer binding efficiency, (5) low read coverage in the primer sequences to avoid targeting loci in repetitive areas of the genome, (6) optimal primer size 20–25 bp, and (7) primers not directly flanking the microsatellite locus. Using these conditions, we selected primer pairs for testing. Primer pairs were synthesized by Eurofins Genomics (Luxembourg).

Primer testing

We attempted to amplify microsatellites in samples of *O. anthropophora*, *O. militaris*, *O. purpurea*, and *O. simia* (Appendix S2: Table S2) and used gel electrophoresis to inform which primer pairs amplified a single product that was polymorphic between species. Pairs of primers that successfully amplified a product were then ordered with a tagged fluorescent label (FAM or JOE). The PCR mix contained 6 μL of 2× DreamTaq PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), 0.5 μL of bovine serum albumin 0.4% (w/v) (BSA), 0.5 μL of labelled forward primer (10×), 0.5 μL of reverse primer, 1 μL of DNA (15 ng/μL), and 1.5 μL of H_2O to make up a final volume of 10 μL. PCR amplifications were carried out using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA). The thermal cycler conditions were as follows: an initial denaturation stage of 3 min at 94°C; followed by 30 cycles of 30 s denaturation at 94°C, 30 s at locus-specific annealing temperatures (Appendix S2: Table S3), and a 45 s of elongation at 72°C; and a final extension of 7 min at 72°C. PCR products were separated on an ABI 3730 sequencer (Applied Biosystems) with 10 μL of HiDi Formamide and 0.15 μL of GeneScan 500 ROX Size Standard (Applied Biosystems). Allele sizes were determined using GeneMapper v. 5 (Applied Biosystems) (Chatterji and Pachter, 2006).

Genetic diversity and population structure

Analyses were first conducted on individual populations and subsequently on data pooled for each species pair (HYB, ANG, and BER). Genetic diversity summary statistics and analysis of molecular variance (AMOVA) were calculated using GENALEX v. 6.5 (Peakall and Smouse, 2006, 2012). Principal component analysis (PCA) was conducted in R v. 3.6.2 using the adegenet package; results were plotted using GGPLOT2 (Wickham, 2009; R Core Team, 2020).

We used thermodynamic integration using the R package rmaverick to detect admixture between the four parental species (Verity and Nichols, 2016). Preliminary analyses were run to identify appropriate parameters. In our final analyses of the hybrid zones, we ran five repetitions of 5000 burn-in steps, 150,000 Markov chain Monte Carlo (MCMC) iterations and 10 rungs using the admixture model. The admixture coefficient (q) generated using rmaverick was used to categorize individuals into parental species and hybrids. We used the R package ggplot2 to display the output from rmaverick graphically.

Admixture analysis and calculation of hybrid indices

Nuclear admixture proportions were estimated using a Bayesian analysis implemented in the program NewHybrids v. 1.0 (Anderson and Thompson, 2002) for individual populations and for pooled data. This model works under the assumption that all admixed samples are early-generation hybrids, the genomes of which are contributed to by two parental gene pools. NewHybrids uses allopatric parental populations to estimate allele frequencies, then calculates the posterior probability that sympatric individuals belong to one of six hybrid classes (parent A, parent B, F1, F2, backcross to parent A and backcross to parent B). After running preliminary tests to establish appropriate parameters, we performed 10 independent repeats of 10,000 burn-in steps and 100,000 MCMC iterations using Jeffrey's priors.

Hybrid indices (HI) estimate the proportion of alleles inherited from each parent species. We used the est.h function in the R package introgress (Gompert and Buerkle, 2010) to calculate hybrid indices for all individuals in sympatric populations. This function calculates a maximum likelihood hybrid index estimate for each individual. Hybrid indices fall between zero and one, which correspond to parental individuals of the two species. Plants with $0.01 < HI < 0.99$ were considered admixed individuals.

Hybrid simulations

To evaluate the power of the developed markers to detect the different hybrid classes we used the hybridize function in the R package adegenet to simulate two generations of hybrids (F1s, F2s and backcrosses; $N = 100$ for each class) in each of the three species combinations, using real allopatric individuals as the parental populations (OA1, OM1, OP1 and OS1; Table 1). We calculated hybrid indices and interspecific heterozygosity for all simulated hybrids for comparison with our observed data. The calc.intersp.het function was used to calculate interspecific heterozygosity for all individuals in the simulated hybrid data set and in the real hybrid zones (Gompert and Buerkle, 2010).

Geometric morphometric analysis of labellum shape

Three labella were sampled at random from the inflorescence of each individual and manually scaled and landmarked using tpsDig v. 2.16 (Rohlf, 2010a, b). Geometric morphometric methods have not previously been used to study petal shape in *Orchis*, so the set of 15 homologous landmarks used in this study (Figure 2) was chosen to capture the main aspects of the petal shape and to complement points that have been used for measurements in traditional morphometric analyses in *Orchis* (Bateman and Farrington, 1987; Cozzolino and Aceto, 1994; Bateman et al., 2008). Cross-validated discriminant function analysis was conducted in MorphoJ (Klingenberg, 2011) to test the strength of taxonomic assignment statistically using labellum shape, given their taxonomic classification using molecular data. Generalized Procrustes and principal component analyses were carried out in the R package geomorph (Adams et al., 2018; R Core Team, 2020) (Appendix S1).

RESULTS

Genetic diversity and differentiation in the four parental *Orchis* species

After a screening process, 11 nuclear microsatellite loci displayed robust amplification and were used in the main study (Appendix S2). To determine whether the 11 loci were able to differentiate between the four *Orchis* taxa, we initially removed all samples from hybrid zones, thereby eliminating potentially admixed individuals. We first used 120 individuals that we collected from four allopatric populations in southern France (OA1, OM1, OP1, and OS1; Table 1).

Overall, 117 alleles were identified across the 11 loci (Appendix S2: Table S3). Genetic differentiation was relatively high between the four *Orchis* species (overall $F_{ST} = 0.34$). Mean heterozygosity was greatest in *O. purpurea* ($H_o = 0.57$) and lowest in *O. anthropophora* ($H_o = 0.42$). AMOVA showed that 56% of the molecular variance is found within individuals, 41% among populations, and 3% among individuals. *Orchis anthropophora* and *O. militaris* have the greatest genetic differentiation ($F_{ST} = 0.29$), and *O. militaris* and *O. purpurea* have the lowest ($F_{ST} = 0.24$). Principal component analysis (PCA) was able to differentiate between the four species using the first three principal components, which captured 50.9% of the variability in the data set (Appendix S2: Figure S1). When all samples from allopatric and sympatric populations were analyzed using rmaverick the optimal number of genetic clusters (K) was found to be four, representing the four species (Appendix S2: Figure S2).

Genetic diversity and differentiation in *Orchis* hybrid zones

For all three species combinations (Figure 1), genetic diversity was consistently higher in the hybrid zones than in the allopatric parental populations (Table 2). The pattern of genetic diversity and admixture was similar in all four HYB populations, and in both ANG populations, so we pooled them into their respective groups to create larger data sets. Overall F_{ST} ranged from 0.18 across HYB allopatric and sympatric populations to 0.20 in ANG and BER populations with their respective allopatric populations, which indicates moderate genetic differentiation. Pairwise F_{ST} values were always higher between the allopatric parental populations than between the pooled hybrid zones and either of the relevant allopatric parental populations (Table 3). The number of private alleles ranged from one (*O. militaris*) to seven (*O. anthropophora*), but when considered in the three species pairs, numbers were higher (Table 4). There were similar numbers of private alleles in each parental species in ANG and BER combinations; in the HYB combination, *O. militaris* had nine more private alleles than *O. purpurea* (Table 4).

Principal component analysis (PCA) was conducted on the different species combinations, and individuals were colored by morphological identification in the field (Figure 3). Most individuals appeared to have been accurately identified based on morphology, but some plants that were collected as parental individuals showed evidence of hybrid influence, whereas some putative hybrids appear to cluster with parental individuals (Figure 3). No individuals identified as one parental species were found to cluster with the other parental species.

In the PCA of HYB populations, some individuals cluster with allopatric *O. militaris*, some with allopatric *O. purpurea*, and many putative hybrids are spread between the two groups (Figure 3A). Some individuals collected as *O. militaris* cluster with the hybrids. There is overlap between *O. militaris* and the hybrids, but *O. purpurea* and the hybrids are more distinct. In ANG populations, PCA (Figure 3B) identifies three distinct clusters that represent

TABLE 2 Genetic diversity in allopatric *Orchis* populations and hybrid zones. N_a = number of alleles, H_e = expected heterozygosity, H_o = observed heterozygosity

Population type	Population code	Polymorphic loci (%)	Mean N_a per locus	H_e	H_o
Hybrid zones	ANG1	100.0	6.46	0.64	0.63
	ANG2	100.0	7.18	0.63	0.58
	BER1	100.0	5.55	0.65	0.55
	HYB1	100.0	7.00	0.61	0.50
	HYB2	100.0	7.00	0.58	0.49
	HYB3	100.0	7.73	0.62	0.49
	HYB4	100.0	8.64	0.63	0.53
Allopatric	OA1	100.0	4.27	0.52	0.42
	OM1	81.8	5.46	0.47	0.46
	OP1	81.8	5.55	0.51	0.57
	OS1	100.0	4.09	0.51	0.49

TABLE 3 Pairwise F_{ST} values between allopatric *Orchis* populations and hybrid zones. Hybrid zones are represented by HYB (*O. militaris*–*O. purpurea*), ANG (*O. purpurea*–*O. simia*), and BER (*O. anthropophora*–*O. simia*)

Population 1	Population 2	Pairwise F_{ST}
O. militaris	*O. purpurea*	0.24
O. purpurea	*O. simia*	0.27
O. anthropophora	*O. simia*	0.26
O. militaris	HYB	0.12
O. purpurea	HYB	0.09
O. purpurea	ANG	0.10
O. simia	ANG	0.11
O. anthropophora	BER	0.12
O. simia	BER	0.08

TABLE 4 Number of private alleles calculated for (A) all four *Orchis* species in allopatric populations and (B–D) for allopatric populations of the two parental taxa in each species combination being studied

Group	Species	No. private alleles
A	*O. anthropophora*	7
	O. militaris	1
	O. purpurea	2
	O. simia	2
B	*O. militaris*	24
	O. purpurea	15
C	*O. purpurea*	29
	O. simia	30
D	*O. anthropophora*	31
	O. simia	29

O. purpurea, *O. simia*, and putative hybrid individuals. Three individuals identified in the field as putative *O. purpurea* backcross hybrids fall closer to the *O. purpurea* cluster than to the hybrid cluster, and one morphological *O. simia* clusters with the putative hybrids, with at least one between the two groups. In the BER population, PCA identifies a cluster of *O. simia*, a cluster of *O. anthropophora*, and a cluster of putative hybrids (Figure 3C). Most individuals fall into one of these three clusters, but a few hybrids lie outside the main hybrid cluster.

Admixture in *Orchis* hybrid zones and hybrid detection

Given that clusters are tools used to summarize the data, rather than definitive statistics, and that each species forms its own cluster (Appendix S2: Figure S2), we have reported rmaverick results based on $K = 2$ clusters, representing the two parental species, as this will be most useful for observing the genomic contribution made by the two parental species and making inferences about hybridization (Figure 4A–C) (van Hengstum et al., 2012; Meirmans, 2015). Allopatric populations have been included at either end as references and were mostly composed of parental individuals, with high average genomic proportions for *O. anthropophora* (0.993), *O. militaris* (0.996), *O. purpurea* (0.994), and *O. simia* (0.990) contributed by their respective genetic clusters. All hybrid zones contained a mixture of hybrid and parental individuals.

In HYB populations the genotypes of both parental taxa influenced the genotypes of plants in the hybrid zones, but on average plants had a greater contribution from the *O. militaris* genome (0.69) than from the *O. purpurea* genome (0.31), despite equal sampling of parental taxa based on morphology (Figure 4A). Parental individuals of both taxa are present in all four sympatric HYB populations. There is a rapid fall in *O. purpurea* q values between 1 and 0.2 before a gradual decline from 0.2 to 0. Only 2.1% of hybrid individuals ($0.01 < q < 0.99$) have mid-range q values between 0.4 and 0.6.

To assess the accuracy of field identification, we used a threshold q value of 0.99 to determine whether plants were parental or admixed individuals. Of the 386 individuals in HYB populations, 319 (82.6%) were accurately identified as parental or admixed individuals in the field using morphology, with 97 identified as *O. purpurea* and 99 as *O. militaris*. However, 43 individuals identified by morphology as *O. militaris* and 13 individuals identified as *O. purpurea* showed evidence of hybrid influence; 11 individuals collected as hybrids had $q > 0.99$ for the *O. militaris* cluster, but no individuals collected as hybrids had $q > 0.99$ for the *O. purpurea* cluster.

In ANG populations, *O. purpurea* (0.57) and *O. simia* (0.43) made similar average genomic contributions to individuals in the hybrid zones (Figure 4B). Examples of parental individuals are found in both hybrid zones, with similar numbers of each. *Orchis purpurea* q values decrease quickly from 1 to 0.7, gradually from 0.7 to 0.4 and then quickly from 0.4 to 0. In total, 37.0% of hybrid individuals ($0.01 < q < 0.99$) have mid-range q values between 0.4 and 0.6.

Seventy-four of 93 (79.6%) sympatric individuals were accurately identified as parental or admixed individuals using morphology in the field, with 14 plants of *O. purpurea* and 10 of *O. simia* identified using the nuclear microsatellites. In total, seven individuals identified by morphology as *O. purpurea* and 11 individuals identified as *O. simia* showed evidence of hybrid influence; two individuals collected as hybrids had $q > 0.99$ for the *O. purpurea* cluster, but no individuals collected as hybrids had $q > 0.99$ for the *O. simia* cluster.

On average, *O. anthropophora* (0.47) and *O. simia* (0.53) made similar contributions to individuals in the BER

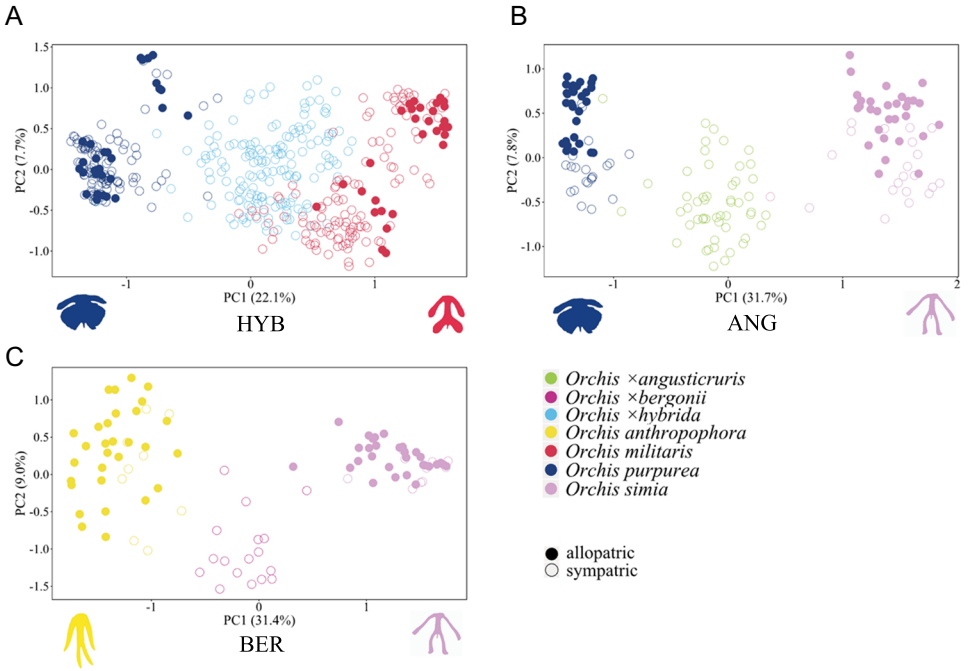

FIGURE 3 Principal component analysis of *Orchis* hybrid zones genotyped using 11 nuclear microsatellite loci. Solid circles represent individuals from allopatric populations and open circles represent individuals from sympatric populations (hybrid zones) (A) HYB, (B) ANG, (C) BER. Individuals are colored according to morphological identification in the field

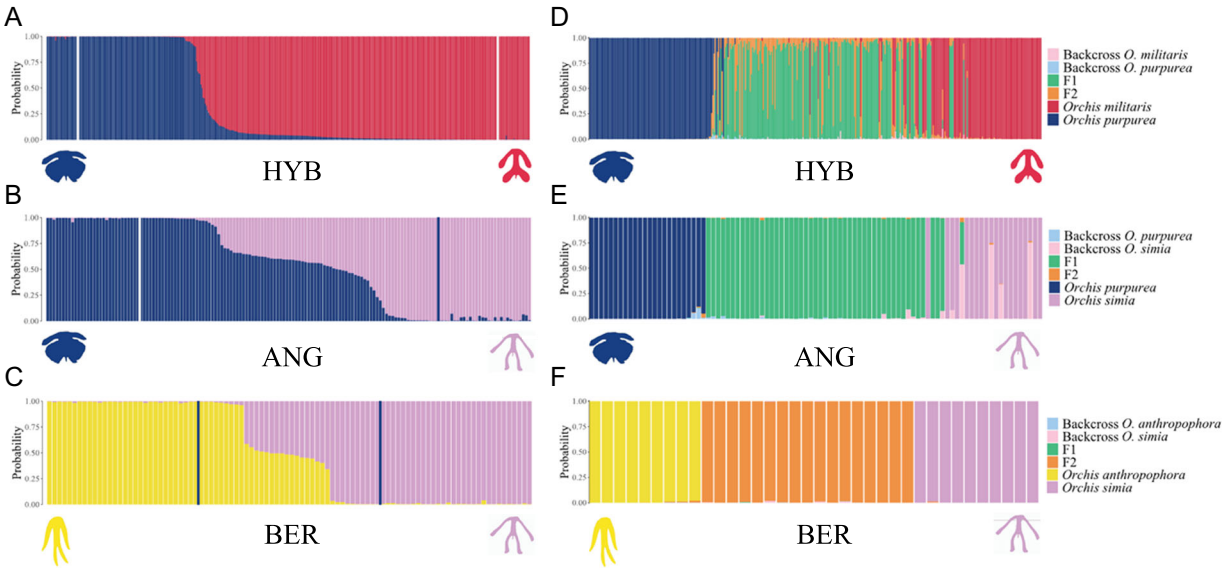

FIGURE 4 (A–C) Clustering admixture analysis of (A) HYB ($N = 128$), (B) ANG ($N = 93$), and (C) BER ($N = 36$) populations using rmaverick. Columns represent individuals and colors represent the genomic proportion assigned to $K = 2$ clusters. Allopatric populations of parental taxa ($N = 30$) have been included at either end, separated by an empty column. Individuals in the hybrid zones are ordered left to right by decreasing genomic contribution from *O. purpurea* in (A) and (B) and from *O. anthropophora* in (C). (D–F) Bayesian inference of hybrid class using NewHybrids. Each column represents an individual in a hybrid zone. Hybrid classes are represented by colors and the extent of each color in a column indicates the posterior probability that an individual belongs to this class. Individuals are in the same order as in the rmaverick plots, but allopatric populations are not included

population (Figure 4C). There are examples of both parental species present. The plants collected as putative hybrids all had intermediate q values (0.4–0.6), and there was little evidence of backcross hybrids in either direction. All individuals were identified correctly as parents or hybrids using morphology in the field.

Assigning parental and hybrid status using NewHybrids

We used NewHybrids software to assign individuals in the hybrid zones to different hybrid classes (Figure 4D–F). Using a threshold q value of 0.8, we found that 349 plants

(94.8%) in the HYB populations could be unequivocally placed in one of four classes (*O. militaris*, *O. purpurea*, F1, and F2). Of these, there were 107 *O. purpurea* individuals, 100 *O. militaris*, 126 F1, and 16 F2 (Figure 4D). No backcrosses to either *O. militaris* or *O. purpurea* were detected. All four HYB populations include parental taxa and F1 and F2 hybrids.

In ANG and BER populations, the number of individuals that NewHybrids could categorize were similar. Using a threshold *q* value of 0.8, we found that 89 plants in ANG populations (95.7%) were successfully placed into one of three classes, with 24 *O. purpurea*, 17 *O. simia*, and 48 F1 hybrids (Figure 4E). In addition, four plants collected as *O. simia* show some evidence of being backcrosses, but they did not reach the threshold *q* value and therefore could not be unequivocally classified. In the BER population, all 36 plants (100.0%) could be confidently categorized, with nine *O. anthropophora*, 10 *O. simia*, and 17 F2 hybrids. F1 hybrids were apparently absent (Figure 4F).

Calculating hybrid indices using R package introgress

Individuals belonging to allopatric populations ($N = 30$ for each) were used as references to calculate a hybrid index (HI) for each individual in the hybrid zones based on allele frequencies. Parental individuals were detected in all populations, whereas most admixed individuals had intermediate hybrid indices indicative of F1 and F2 hybrids (Figure 5). Lower and higher hybrid indices provide some evidence of backcross hybrids.

In HYB populations, hybrid indices ranged from 0 (*O. purpurea*) to 1 (*O. militaris*). In total, 100 plants had a hybrid index <0.01, whereas 26 plants had a hybrid index >0.99, suggesting there were considerably fewer individuals of *O. militaris* than of *O. purpurea* (Figure 5A). Hybrid indices rose sharply from 0 to 0.4, followed by a gradual increase from 0.4 to 1 (Figure 5A). Only 83 plants (21.5%) had a hybrid index between 0.01 and 0.50, whereas 177 plants (45.9%) had a hybrid index between 0.50 and 0.99, suggesting most hybrid individuals ($0.01 < HI < 0.99$) had a greater proportion of *O. militaris* DNA in their genomes. In total, 70.4% of individuals identified using NewHybrids as F1 or F2 hybrids were assigned hybrid indices between 0.4 and 0.6. Furthermore, 92.5% of *O. purpurea* as identified by NewHybrids had a hybrid index of <0.01, but only 26.0% of individuals identified as *O. militaris* had a hybrid index >0.99, with 59.0% having hybrid indices ranging from 0.7 to 0.99 (Figure 5A).

When considered as separate populations, hybrid indices suggest all four hybrid zones had many *O. purpurea* individuals, but that only populations HYB1 and HYB2 had *O. militaris* individuals (Appendix S2, Figure S3). HYB3 and HYB4 had few, if any, *O. militaris*. All four populations have individuals with hybrid indices indicative of *O. militaris* backcrosses. HYB4 also appeared to have many individuals that could constitute backcrosses to *O. purpurea*. These were rare in the other three HYB populations.

In ANG populations, where 0 represents *O. purpurea* and 1 represents *O. simia*, there are 18 individuals with HI <0.01 and seven individuals with HI >0.99 (Figure 5B), implying there were more *O. purpurea* than *O. simia*, despite equal sampling based on morphology. There were 47 plants with a hybrid index between 0.01 and 0.50 (50.5%) and 24 plants between 0.50 and 0.99 (25.8%) (Figure 5B), suggesting admixed individuals had a greater proportion of *O. purpurea* DNA in their genomes. There was a rapid increase in hybrid indices between 0.1 and 0.3, a gradual increase from 0.3 to 0.6 and then another rapid increase between 0.6 and 0.9.

The pattern of hybridization was similar in both ANG populations when considered separately (Appendix S2: Figure S3). *Orchis purpurea* was present in both populations, but *O. simia* was rare or absent in ANG1 (though there were many more individuals identified by morphology as *O. simia* present in this population than there were of putative hybrids or *O. purpurea*). Both populations had individuals with hybrid indices indicative of backcrosses, particularly in ANG2, but all except three of these were classified as parental individuals by NewHybrids. The three plants that could not be confidently assigned a hybrid class by NewHybrids were most likely to be backcrosses to *O. simia* (Figure 4E). One anomalous individual identified using NewHybrids and morphology as *O. simia* had a hybrid index of 0.57.

In the BER population, 0 represents *O. anthropophora*, and 1 represents *O. simia* (Figure 5C). Three plants had a hybrid index <0.01, and eight plants had a hybrid index >0.99. There were sharp increases in hybrid index between 0.2 and 0.4 and between 0.7 and 0.9, but there were plants with hybrid indices, suggesting they are not parental individuals and could constitute backcrosses. These were all identified as parental individuals by NewHybrids and using morphology in the field. The putative hybrids, identified as F2s by NewHybrids, all had intermediate hybrid indices.

Comparing hybrid zones to simulated F1, F2, and backcross hybrids

PCA, rmaverick analysis, and introgress hybrid index estimates all suggested there were many backcrossed individuals present, particularly in HYB populations. However, analysis using NewHybrids failed to detect any backcrossed individuals. Given the apparent discrepancy, we wanted to test the ability of NewHybrids and INTROGRESS to distinguish between the different hybrid classes. We analyzed two more allopatric *O. militaris* (OM2) and *O. purpurea* (OP2) populations and found that NewHybrids unequivocally assigned parental status to all plants; all individuals in OM2 had a hybrid index >0.99, and all individuals in OP2 had a hybrid index <0.01 (data not shown). NewHybrids failed to identify any backcrossed individuals that other analyses had

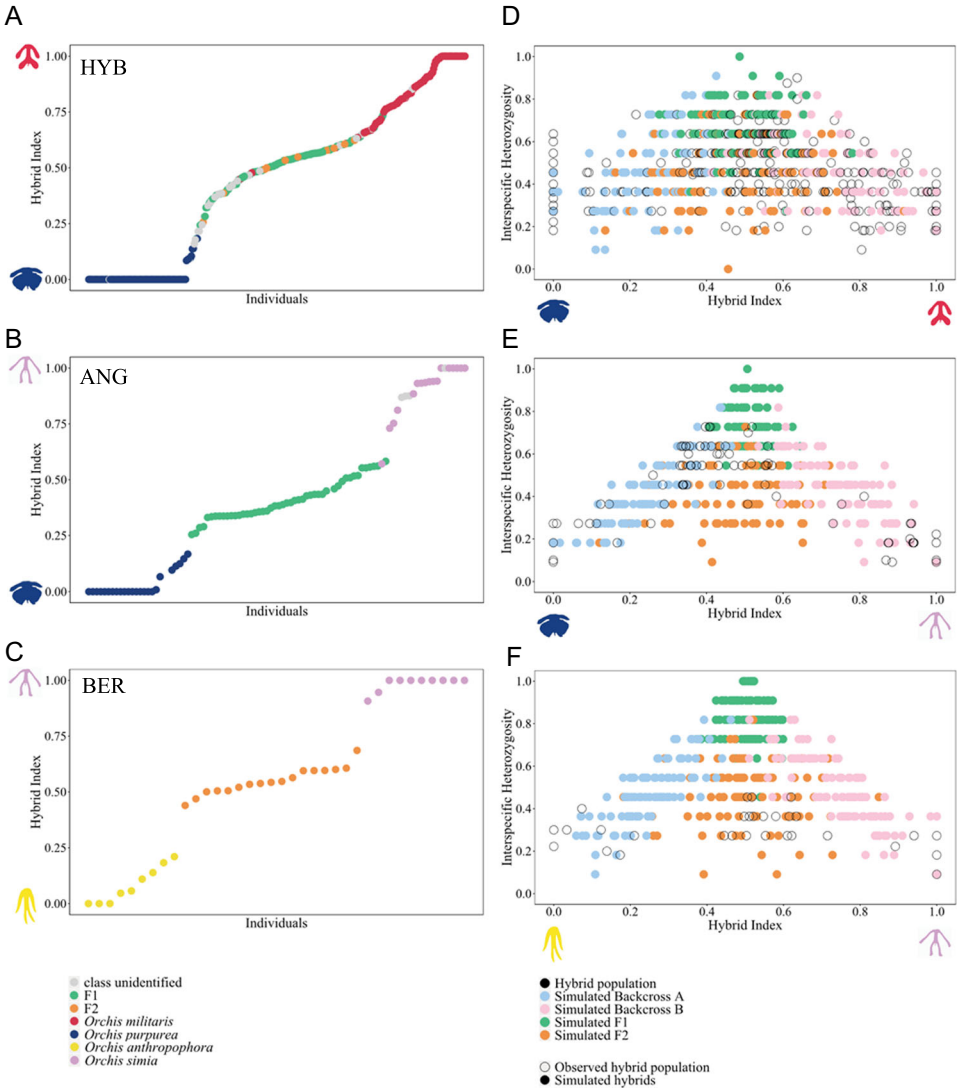

FIGURE 5 (A–C) Hybrid indices calculated using the R package introgress for individuals in (A) HYB, (B) ANG, and (C) BER populations. A hybrid index provides the fraction of the genome inherited from *O. militaris* (A) or *O. simia* (B, C). Individuals are colored by the hybrid class assigned by NewHybrids in Figure 4. (D–F) Summary of interspecific heterozygosity plotted against hybrid index for plants in observed hybrid zones (open circles) and simulated hybrids (closed circles). Simulated backcross A is backcrossed to *O. purpurea* (A, B) or *O. anthropophora* (C); simulated backcross B is backcrossed to *O. militaris* (A) or *O. simia* (B, C)

unanimously concluded were present which implied NewHybrids was failing to pick up backcross genotypes present in the sympatric populations.

To investigate this further, we analyzed simulated hybrids using NewHybrids and found that simulated ANG and BER hybrids were classified with greater confidence than HYB simulated hybrids (Appendix S2: Table S4). F1s were most successfully identified in all three cases. In total, 80.5% and 79.8% of simulated hybrids met the threshold posterior probability value of 0.8 in ANG and BER combinations, respectively. However, when simulated HYB hybrids were analyzed only 35.3% of individuals met the 0.8 threshold, including only 21% of simulated backcrosses. Only 16% of simulated *O. militaris* backcrosses were confidently identified.

As expected, interspecific heterozygosity was consistently highest in simulated F1 hybrids and lower in simulated F2 and backcross hybrids (Figure 5D–F). In simulated backcross hybrids interspecific heterozygosity decreased with decreasing or increasing hybrid indices, depending on the direction of hybridization. Also as expected, hybrid indices were intermediate for simulated F1 and F2 hybrids, with the F2s exhibiting a greater range of values. Simulated hybrids in ANG and BER combinations exhibited similar, relatively clear patterns. This pattern was also observed in simulated HYB hybrids, although the different hybrid classes were less distinct.

When comparing interspecific heterozygosity in our observed data with the simulated results, we found that plants in HYB populations most closely resembled simulated F1s, F2s, and backcrosses (Figure 5D). In ANG populations, most observed individuals fell between simulated F1 and simulated *O. purpurea* backcrosses, but several observed plants resembled simulated *O. simia* backcrosses and

F2s (Figure 5E). Plants observed in the BER population best fit the F2 simulated cases, but there were some that resembled simulated backcrosses (Figure 5F).

Morphological analyses

Discriminant analysis (Figure 6) and principal component analysis (Appendix S2: Figures S4 and S5) revealed that the labellum shape of hybrids in HYB populations exhibited a small overlap with that of *O. purpurea* and a large overlap with that of *O. militaris*. In these pairs of taxa, *O. militaris* and *O. purpurea* were more accurately classified based on labellum shape than the hybrids (Appendix S2: Table S5). Of 249 admixed individuals, ca. 72–89% were correctly classified as hybrids using labellum shape.

Hybrids in ANG and BER populations were more readily distinguished from their parental species, though some overlap was observed (Figure 6; Appendix S2: Figure S5). Hybrids were correctly classified >95% of the time in ANG populations (though 12 individuals, including some identified as potential backcrosses, were missing morphological data), and in the BER population they were correctly classified ca. 81–87% of the time (Figure 6; Appendix S2: Table S5). When combining shape data with nuclear microsatellite data, there was a strong positive correlation between molecular and morphological hybrid indices in HYB ($R = 0.82$, $P < 0.001$), ANG ($R = 0.86$, $P < 0.001$), and BER ($R = 0.79$, $P < 0.001$) populations (Figure 7).

DISCUSSION

We used novel nuclear microsatellite markers and geometric morphometrics to show that genetic admixture has occurred in *O. militaris*–*O. purpurea*, *O. purpurea*–*O. simia*, and *O. anthropophora*–*O. simia* hybrid zones in southern France. Our data suggest that hybridization has gone beyond the first generation of hybrids in all three combinations. They show formerly undocumented patterns of hybridization in *O. purpurea*–*O. simia* and *O. anthropophora*–*O. simia* hybrid zones and corroborate, across multiple French populations, the extensive hybridization previously reported from an *O. militaris*–*O. purpurea* hybrid zone in Belgium (Jacquemyn et al., 2012a).

When different species grow together and flower at the same time, species boundaries can be maintained by attracting different pollinators or by post-zygotic reproductive barriers (Cozzolino et al., 2006). These four *Orchis* species have the same number of chromosomes ($2n = 42$) (Cozzolino et al., 2004; Kretzschmar et al., 2007) and have overlapping phenology, pollinator communities, and habitat preferences (Cozzolino et al., 2004; Kretzschmar et al., 2007; Schatz et al., 2020). As a result, it is likely that pollen is sporadically transferred between the different species when growing in sympatry.

In all three studied species combinations, we observed hybrid zones containing both parental and admixed individuals, suggesting that the degree of reproductive isolation is strong enough for the parental species to endure over time despite growing in sympatry. It is unlikely that all the studied populations have come so recently into sympatry that only a few generations of interspecific hybridization have elapsed, and certainly the sympatric populations analyzed do not represent recent hybrid zones between large parental populations. Interspecific hybridization may be less common than intraspecific reproduction in these hybrid zones. However, given the ease with which these species hybridize, the presence of apparently "pure" parental individuals in hybrid zones is more likely to be because post-zygotic barriers limit reproduction between hybrids and parental individuals.

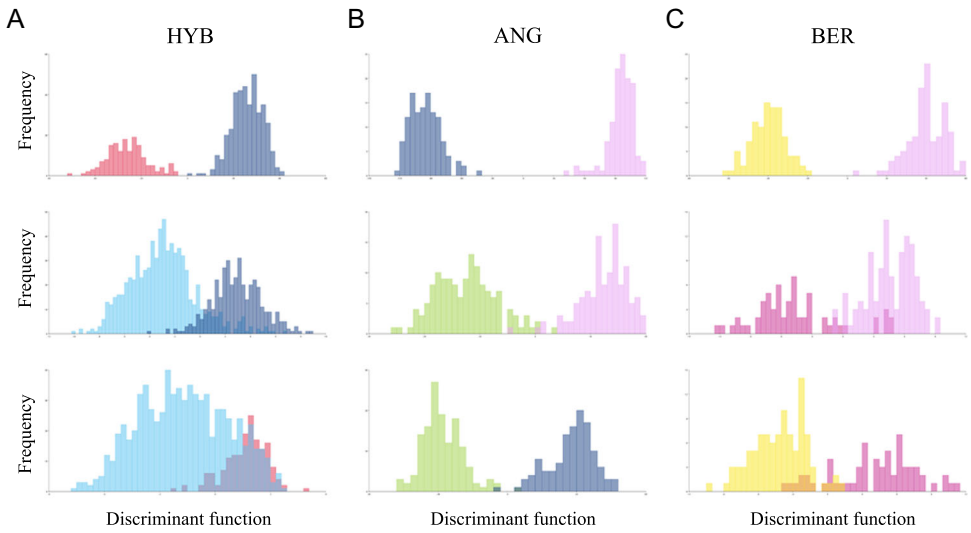

FIGURE 6 Discriminant analysis of the shape of labella in (A) HYB, (B) ANG, and (C) BER populations with *O. anthropophora* (yellow), *O. militaris* (red), *O. purpurea* (navy), *O. simia* (lilac), *O. ×hybrida* (light blue), *O. ×angusticruris* (green), and *O. ×bergonii* (magenta). Individuals were assigned taxonomic status using nuclear microsatellite data

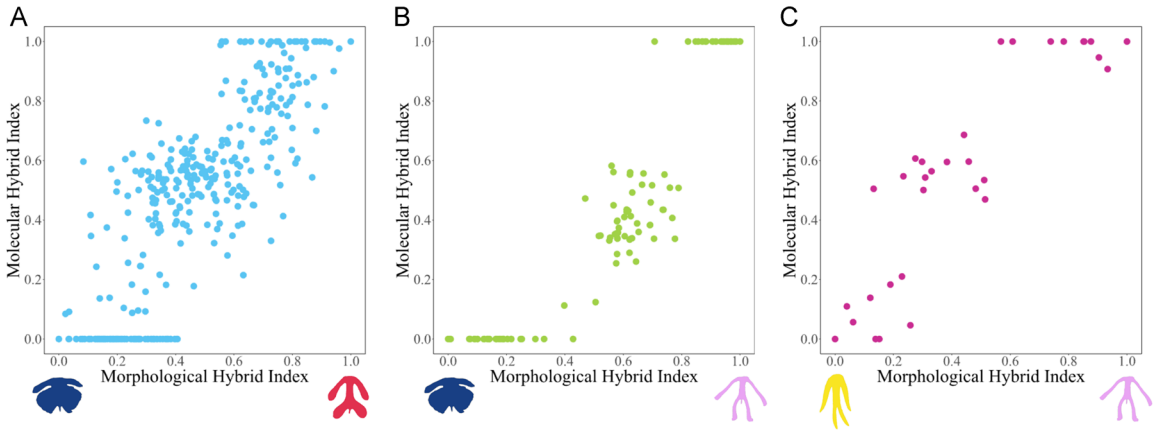

FIGURE 7 Correlation between molecular hybrid index (calculated using introgress) and morphological hybrid index (calculated using the first principal component from PCA of Procrustes shape [Appendix S5]) in *Orchis* hybrid zones: (A) HYB, (B) ANG, and (C) BER populations. Hybrid indices of 1 are indicative of *O. militaris* (A) or *O. simia* (B, C). Representative labella indicate which taxa the limits represent

The co-existence of parental and hybrid individuals in these *Orchis* hybrid zones implies that these populations are tension zones (Barton and Hewitt, 1985), being maintained by a balance between gene flow into the hybrid zone and selection acting against admixed individuals (Arnold, 1999; Abbott and Brennan, 2014). Such a balance has been observed in *Populus* and *Senecio* hybrid zones (Martinsen et al., 2001; Brennan et al., 2009). Under the tension zone model, admixed individuals might have reduced fecundity relative to parental forms due to epistatic incompatibilities between genes inherited from the parental species (Abbott and Brennan, 2014). This, coupled with the reduced frequency of interspecific hybridization relative to intraspecific reproduction, may be sufficient to maintain parental types.

Asymmetric backcrossing in multiple *O. militaris–O. purpurea* hybrid zones

Among the four *Orchis* species studied here, *O. militaris* and *O. purpurea* are widely considered to be the most closely related (Bateman et al., 2003; Jacquemyn et al., 2011b; L. Bersweden et al., unpublished data). Although there are obvious morphological differences, reproductive isolation between these taxa is weak (Jacquemyn et al., 2012b). Experimental crosses conducted between *O. militaris* and *O. purpurea* found that both have high fruit set (>75%) when pollinated with pollen taken from the other species (Jacquemyn et al., 2011b). Reproductive isolation, calculated as a function of both pre- and post-zygotic barriers, is reported to be 0.55 on average, which is much lower than between these two species and *O. anthropophora* (0.79 and 0.88, respectively) (Jacquemyn et al., 2012b).

Our results showed that allopatric populations exhibit relatively well-defined genetic and morphological clusters for each species (Appendix S2: Figures S1 and S5A) and analyses of hybrid zones consistently suggested that genetic admixture has occurred between these two species. Additionally, our data revealed that at least two generations have passed since hybridization first occurred, a finding that is consistent with other results from a hybrid zone in Belgium and observations in the field (Kretzschmar et al., 2007; Jacquemyn et al., 2012a).

Hybrids formed a range of intermediate labellum phenotypes between the parental taxa rather than forming a morphologically homogeneous group. Molecular results from rmaverick and hybrid index estimations from introgress, combined with geometric morphometric data, suggest that hybridization was bidirectional but that backcrossing was asymmetric toward *O. militaris*, which typically made a greater contribution to the genomic composition of the hybrid zone. *Orchis purpurea* backcrosses were present but rare.

This asymmetry can be explained by considering the flowering phenology of the two species in question. In general, *O. purpurea* starts to flower 7–10 days earlier than the hybrids and 10–14 days earlier than *O. militaris* (Kretzschmar et al., 2007). Potentially, *O. purpurea* stigmas are less likely to be receptive to pollen at the onset of flowering of the hybrids (either because they have already been pollinated or because unfertilized stigmas have begun to deteriorate during the later stages of anthesis); nevertheless, there is some pollination from hybrids, giving rise to some *O. purpurea* backcrosses. Hybrids can receive pollen from other hybrids, *O. militaris* or *O. purpurea*, but the relative rarity of hybrids reduces the chance of hybrid–hybrid pollination, so there are only a few F2 individuals. The evidence suggests that *O. militaris* plants are more readily fertilized by hybrid pollen than are *O. purpurea* plants, perhaps because the receptiveness of *O. militaris* stigmas overlaps more with hybrid pollen production than it does for *O. purpurea*.

The data could also be explained by the longevity of *Orchis* pollen, potentially meaning that the earlier-flowering species is more likely to be the pollen donor in interspecific hybridization events (Neiland and Wilcock, 1995; Bellusci

et al., 2010). The duration of pollen viability is 20–30 days for *O. anthropophora*, *O. italica*, *O. mascula*, and *O. provincialis* (Bellusci et al., 2010), suggesting viability of 20–30 days is also likely for *O. purpurea*. If *O. purpurea* pollen is still viable when *O. militaris* begins to flower, then successful fertilization of *O. militaris* plants with *O. purpurea* pollen will be considerably more likely than the reverse. *Orchis militaris* would then be the female parent more often than the male parent, so we might expect the spatial distribution of hybrid plants to display greater overlap with *O. militaris* plants than with *O. purpurea*, potentially contributing to pollen flow between hybrids and *O. militaris*, which could lead to the observed asymmetry in these populations (Appendix S2: Figure S3).

The apparent asymmetry toward *O. militaris* contrasts with findings from an *O. militaris–O. purpurea* hybrid zone studied in Belgium using AFLP markers (Jacquemyn et al., 2012a). The authors found that hybridization was bidirectional but asymmetric in favor of *O. purpurea* rather than *O. militaris*. Heterogeneous patterns of hybridization between the same species such as this are not uncommon (Zeng et al., 2011; Zanella et al., 2016) and can be a result of differences in the relative abundance of each species in different geographic locations (Lepais et al., 2009) or local ecological factors that impact the way in which selection acts on the hybrids (Williams et al., 2001; Michalski and Durka, 2015). The heterogeneity observed in geographically distant *O. militaris–O. purpurea* hybrid zones could be due to a range of different factors (e.g., differential association with mycorrhizal fungi or variable overlap in flowering phenology at different latitudinal gradients). Greater sampling across their shared geographic range is required to establish the mechanisms driving this pattern.

F1 dominance in *O. purpurea–O. simia* hybrid zones despite hybrid fertility

Our molecular and morphological data indicated that putative *O.* ×*angusticruris* plants are natural hybrids between *O. purpurea* and *O. simia*. Molecular data revealed that backcrossing had occurred in both directions, but that populations were largely composed of F1 hybrids. Hybrid zones dominated by fertile F1s have been observed in *Rhododendron* and *Alnus* (Milne et al., 2003; Šmíd et al., 2020). Reduced extrinsic fitness relative to F1s in specific habitats resulted in post-F1 hybrids being outcompeted (Milne et al., 2003; Šmíd et al., 2020). However, given the similarity in habitat preferences exhibited by *O. purpurea* and *O. simia*, the relative rarity of post-F1 hybrids is perhaps more likely a result of lower intrinsic fitness relative to F1s, potentially caused by genetic incompatibilities and/or recombination breaking up selectively favorable allelic combinations, as is the case in some European *Populus* hybrid zones (Christe et al., 2016).

Absence of F1 hybrids in an *O. anthropophora–O. simia* hybrid zone

A number of morphological analyses (Bateman and Farrington, 1987; Cozzolino and Aceto, 1994), preliminary data from ribosomal ITS sequences (R. Smith, B. Schatz, and M. F. Fay, unpublished data) and field observations all suggest that hybrids between *O. anthropophora* and *O. simia* are mostly limited to the F1 generation. It was thought that genetic distance between *O. anthropophora* and the other three species meant hybridization was unlikely to occur beyond the first generation (Cozzolino et al., 2004). Fruit set in *O.* ×*bergonii* is 10 times lower than in the parental taxa, which may act as part of a series of reproductive barriers to prevent gene flow (Schatz, 2006; Schatz et al., 2010; B. Schatz, unpublished data). The hybrid has also been shown to emit larger quantities and different proportions of volatile compounds than either parent (Schatz et al., 2010). Certain compounds reaching higher concentrations in the hybrid could act to repel pollinators. Furthermore, some compounds known to be involved in orchid pollinator attraction (e.g., some benzenoids) were found to be present in significantly lower quantities in *O.* ×*bergonii* than in the parental taxa and other orchid hybrids (Knudsen et al., 2006; Salzmann et al., 2007; Schatz et al., 2010), which implies that hybridization beyond the first generation is unlikely to occur.

Our results, however, showed that F1 individuals are absent from the studied hybrid zone and that hybrids with intermediate morphology were F2s rather than F1s. This finding in itself is not unusual as there are many examples of studies on the structure of hybrid zones formed by different hybridizing plant species in which first-generation hybrids are rare or absent (e.g., Arnold et al., 2010; Zeng et al., 2011). If cross-pollination events between hybrids are rare in *O. anthropophora–O. simia* hybrid zones, it is plausible that in long-standing populations the initial F1 generation is outlived by the subsequent F2 generation, resulting in the presence of F2 hybrids and the absence of F1s. A larger sample size would be needed from multiple hybrid zones to determine whether F1 individuals are truly rare or whether the studied population is the exception.

If F1 hybrids are fertile, then backcross hybridization also could have occurred at some point in the past. There was some suggestion from molecular and morphological hybrid indices that admixture with parental individuals might have occurred; these individuals could potentially be later-generation backcrosses that are not detected by NewHybrids and are the descendants from an historical backcrossing event. Given that the hybrids studied here are F2 individuals, we can infer that F1 hybrids are fertile, and therefore backcrossing in either direction remains a possibility, so historical gene flow between the two species cannot be ruled out. While distinguishing between *O. anthropophora* and *O. simia* using morphology in the field is a simple task, subtle information about hybridization such as this is more difficult for humans to detect. Geometric

morphometrics could prove a useful tool in identifying the presence of potential late-generation backcrosses without having to genotype individuals when making informed decisions about conservation management and sampling sites in future studies.

Variation in the frequency and extent of hybridization

Hybrids can be common when parental taxa are found growing in sympatry. However, our results imply that the frequency and extent of contemporary hybridization vary among species combinations. Firstly, *O. purpurea* hybridizes more frequently with *O. militaris* than it does with *O. simia*, perhaps because they grow together more frequently. In southern France, *O. purpurea* usually prefers lower elevations, whereas *O. simia* typically grows at higher elevations. *Orchis militaris* is more often found growing with *O. purpurea* than either species is with *O. simia* (Kretzschmar et al., 2007). Alternatively, morphological data presented here shows that the legs and torso of *O. militaris* and *O. purpurea* labella are wider than those of *O. simia*. These differences might be driven by pollinator preference among the largely overlapping collection of generalist pollinators that visit these orchids. For example, larger insects like bumblebees might have more difficulty landing on the labella of *O. simia* as they have narrow appendages, but the wider torso and legs of *O. militaris* and *O. purpurea* labella may facilitate landing, therefore making hybridization between these two species more likely.

Furthermore, where hybrids are observed, the number of individuals and the frequency of backcrosses appear much greater in *O. militaris–O. purpurea* populations. Thus, the disparity in hybrid frequency is perhaps more likely to be a consequence of genetic distance, as more closely related diploid taxa are more likely to hybridize and produce homoploid hybrids (Montalvo and Ellstrand, 2001; Edmands, 2002; Mallet, 2005; Paun et al., 2009). *Orchis militaris* and *O. purpurea* are considered sister species, with *O. simia* being more distantly related (Bateman et al., 2003). It is likely that there will be fewer genetic incompatibilities between *O. militaris* and *O. purpurea*, and therefore intrinsic selection acting on the hybrids will be weaker than in *O. purpurea–O. simia* hybrid zones. As a result, hybridization between *O. militaris* and *O. purpurea* is likely to be more extensive, in the absence of other selective pressures.

CONCLUSIONS

The 11 newly isolated microsatellites in this study, alongside the set of loci for *Orchis* subgenus *Masculae* developed in parallel by Calevo et al. (2021), are the first nuclear microsatellite markers characterized in the genus *Orchis*. Our molecular and morphological data collectively indicate that contemporary hybridization has occurred in *O. militaris–O. purpurea*, *O. purpurea–O. simia*, and *O. anthropophora–O. simia* hybrid zones, that F1 hybrids are fertile and that at least two generations have passed since the initial hybridization events took place. Hybridization in the studied *O. purpurea–O. simia* and *O. anthropophora–O. simia* hybrid zones was largely limited to early generation hybrids, but further admixture had occurred. The *O. militaris–O. purpurea* hybrid zones exhibited extensive admixture with backcrossing asymmetric towards *O. militaris*. These patterns were reflected in labellum geometric morphometric data, which correlated strongly with nuclear microsatellite data in all three species combinations.

ACKNOWLEDGMENTS

We thank László Csiba, Penny Malakasi, and Dion Devey (all RBG, Kew) for help with sequencing and technical support. This work was supported by NERC, the Bentham-Moxon Trust, the Systematics Research Fund, the Botanical Research Fund, and the OSU-OREME. We are also very grateful to the reviewers and editors whose comments made a big difference to the quality of this study.

AUTHOR CONTRIBUTIONS

Ideas and experimental design were conceived by L.B., M.F.F., A.R.L., and J.J.C. Fieldwork was carried out by L.B. and B.S. Primer design was conducted by L.B., J.C., and J.V. Genome skimming was conducted by L.B., R.G., and R.C. Microsatellite genotyping was carried out by L.B. and J.H. Microsatellite data analysis was carried out by L.B., J.V., and A.J. Geometric morphometric data was collected and analyzed by L.B. The manuscript was written by L.B. and revised by all authors.

DATA AVAILABILITY STATEMENT

Raw microsatellite data are available from the Data Dryad Repository at https://doi.org/10.5061/dryad.p5hqbzkp4 (Bersweden et al., 2021).

ORCID

Leif Bersweden http://orcid.org/0000-0001-6650-4402
Juan Viruel http://orcid.org/0000-0001-5658-8411
Bertrand Schatz http://orcid.org/0000-0003-0135-8154
Roberta Gargiulo http://orcid.org/0000-0001-8663-6568
Jacopo Calevo http://orcid.org/0000-0002-1717-2365
Ana Juan http://orcid.org/0000-0002-2929-3818
James J. Clarkson http://orcid.org/0000-0002-1780-5599
Andrew R. Leitch http://orcid.org/0000-0001-8574-302X
Michael F. Fay http://orcid.org/0000-0003-3491-9093

REFERENCES

Abbott, R. J. 2017. Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *Journal of Systematics and Evolution* 55: 238–258.

Abbott, R. J., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. Boughman, et al. 2013. Hybridization and speciation. *Journal of Evolutionary Biology* 26: 229–246.

Abbott, R. J., and A. C. Brennan. 2014. Altitudinal gradients, plant hybrid zones and evolutionary novelty. *Philosophical Transactions of the Royal Society, B, Biological Sciences* 369: 20130346.

Aceto, S., P. Caputo, L. Gaudio, R. Nazzaro, and S. Cozzolino. 2000. Molecular approach to the identification and characterization of natural hybrids between *Orchis pauciflora* Ten. and *Orchis quadripunctata* Cyr. ex Ten. (Orchidaceae). *Botanica Helvetica* 110: 31–39.

Adams, D. C., M. L. Collyer & A. Kaliontzopou 2018. geomorph: software for geometric morphometric analyses. R package version 3.0.6. Website: https://cran.r-project.org/web/packages/geomorph/index.html

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160: 1217–1229.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Computer program and documentation distributed by Babraham Bioinformatics. Website: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arnold, M. L. 1999. Natural hybridization and evolution. Oxford University Press, Oxford, UK.

Arnold, M. L., S. Tang, S. J. Knapp, and N. H. Martin. 2010. Asymmetric introgressive hybridization among Louisiana *Iris* species. *Genes* 1: 9–22.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* 16: 113–148.

Bateman, R. M., and O. S. Farrington. 1987. A morphometric study of ×*Orchiaceras bergonii* (Nanteuil) Camus and its parents (*Aceras anthropophorum* (L.) Aiton f. and *Orchis simia* Lamarck) in Kent. *Botanical Journal of the Linnean Society* 16: 397–407.

Bateman, R. M., P. Hollingsworth, J. Preston, Y.-B. Luo, A. Pridgeon, and M. W. Chase. 2003. Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). *Botanical Journal of the Linnean Society* 142: 1–40.

Bateman, R. M., R. J. Smith, and M. F. Fay. 2008. Morphometric and population genetic analyses elucidate the origin, evolutionary significance and conservation implications of *Orchis* ×*angusticruris* (*O. purpurea* × *O. simia*), a hybrid orchid new to Britain. *Botanical Journal of the Linnean Society* 157: 687–711.

Bellusci, F., A. Musacchio, R. Stabile, and G. Pellegrino. 2010. Differences in pollen viability in relation to different deceptive pollination strategies in Mediterranean orchids. *Annals of Botany* 106: 769–774.

Bersweden, L., J. Viruel, B. Schatz, J. Harland, R. Gargiulo, R. S. Cowan, J. Calevo, et al. 2021. Data from: Microsatellites and petal morphology reveal new patterns of admixture in Orchis hybrid zones. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.p5hqbzkp4

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Brennan, A. C., J. R. Bridle, A. L. Wang, S. J. Hiscock, and R. J. Abbott. 2009. Adaptation and selection in the *Senecio* (Asteraceae) hybrid zone on Mount Etna, Sicily. *New Phytologist* 183: 702–717.

Calevo, J., R. Gargiulo, L. Bersweden, J. Viruel, C. González-Montelongo, K. Rhebbas, L. Boutabia, and M. F. Fay. 2021. Molecular evidence of species- and subspecies-level distinctions in the rare *Orchis patens s.l.* and implications for conservation. *Biodiversity and Conservation* 30: 1293–1314.

Carrió, E., and J. Güemes. 2014. The effectiveness of pre- and post-zygotic barriers in avoiding hybridization between two snapdragons (*Antirrhinum* L.: Plantaginaceae). *Botanical Journal of the Linnean Society* 176: 159–172.

Chase, M. W., and H. G. Hills. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* 40: 215–220.

Chatterji, S., and L. Pachter. 2006. Reference based annotation with GeneMapper. *Genome Biology* 7: R29.

Christe, C., K. N. Stölting, L. Bresadola, B. Fussi, B. Heinze, D. Wegmann, and C. Lexer. 2016. Selection against recombinant hybrids maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and recurrent gene flow. *Molecular Ecology* 25: 2482–2498.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer, Sunderland, MA, USA.

Cozzolino, S., and S. Aceto. 1994. Morphological and molecular characterization of ×*Orchiaceras bergonii* (Nanteuil) E.G. Cam. *Giornale Botanico Italiano* 128: 861–867.

Cozzolino, S., S. D'Emerico, and A. Widmer. 2004. Evidence for reproductive isolate selection in Mediterranean orchids: karyotype differences compensate for the lack of pollinator specificity. *Proceedings of the Royal Society, B, Biological Sciences* 271: S259–S261.

Cozzolino, S., and A. Widmer. 2005. Orchid diversity: an evolutionary consequence of deception? *Trends in Ecology and Evolution* 20: 487–494.

Cozzolino, S., A. Nardella, S. Impagliazzo, A. Widmer, and C. Lexer. 2006. Hybridization and conservation of Mediterranean orchids: Should we protect the orchid hybrids or the orchid hybrid zones? *Biological Conservation* 129: 14–23.

Edmands, S. 2002. Does parental divergence predict reproductive compatibility? *Trends in Ecology & Evolution* 17: 520–527.

Evans, A., S. Janssens, and H. Jacquemyn. 2020. Impact of climate change on the distribution of four closely related *Orchis* (Orchidaceae) species. *Diversity* 12: 312.

Dryden, I. L., and K. V. Mardia. 2016. Statistical shape analysis: with applications in R, vol. 995. John Wiley, Chichester, UK.

Faircloth, B. C. 2008. msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.

Farrell, L. 1985. Biological flora of the British Isles: *Orchis militaris*. *Journal of Ecology* 73: 1041–1053.

Fay, M. F., R. J. Smith, K. Zuiderduin, E. Hooper, R. Samuel, R. M. Bateman, and M. W. Chase. 2007. How does hybridisation influence the decision making process in conservation? The genus *Orchis* (Orchidaceae) as a case history. *Lankesteriana* 7: 135–137.

Fay, M. F., and I. Taylor. 2015. Orchis anthropophora. *Curtis's Botanical Magazine* 32: 63–71.

Furches, M. S., R. L. Small, A. Furches. 2013. Hybridization leads to interspecific gene flow in *Sarracenia* (Sarraceniaceae). *American Journal of Botany* 100: 2085–2091.

Gompert, Z., and C. A. Buerkle. 2010. introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources* 10: 378–384.

Harrap, A., and S. Harrap. 2009. Orchids of Britain & Ireland: a field and site guide. A & C Black, London, UK.

Henneresse, T., and D. Tyteca. 2016. Insect visitors and potential pollinators of *Orchis militaris* (Orchidaceae) in southern Belgium. *Journal of Insect Science* 16: 104.

Heuertz, M., J. Hausman, O. Hardy, G. Vendramin, N. Frascaria-Lacoste, and X. Vekemans. 2004. Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution* 58: 976–988.

Ilves, A., M. Metsare, K. Tali, and T. Kull. 2015. The impact of recent colonization on the genetic diversity and fine-scale genetic structure in *Orchis militaris* (L.). *Plant Systematics and Evolution* 301: 1875–1886.

Jacquemyn, H., R. Brys, B. P. A. Cammue, O. Honnay, and B. Lievens. 2011b. Mycorrhizal associations and reproductive isolation in three closely related *Orchis* species. *Annals of Botany* 107: 347–356.

Jacquemyn, H., R. Brys, O. Honnay, and I. Roldán-Ruiz. 2012a. Asymmetric gene introgression in two closely related *Orchis* species: evidence from morphometric and genetic analyses. *BMC Evolutionary Biology* 12: 178.

Jacquemyn, H., R. Brys, O. Honnay, I. Roldán-Ruiz, B. Lievens, and T. Wiegand. 2012b. Nonrandom spatial structuring of orchids in a hybrid zone of three *Orchis* species. *New Phytologist* 193: 454–464.

Jacquemyn, H., R. Brys, and M. J. Hutchings. 2011a. Biological flora of the British Isles: *Orchis anthropophora* (L.) All. (*Aceras anthropophorum* (L.) W.T. Aiton). *Journal of Ecology* 99: 1551–1565.

Jacquemyn, H., R. Brys, K. Vandepitte, O. Honnay, and I. Roldán-Ruiz. 2006. Fine-scale genetic structure of life history stages in the food-deceptive orchid *Orchis purpurea*. *Molecular Ecology* 15: 2801–2808.

Joffard, N., F. Massol, M. Grenié, C. Montgelard, and B. Schatz. 2019. Effect of pollination strategy, phylogeny and distribution on pollination niches of Euro-Mediterranean orchids. *Journal of Ecology* 107: 478–490.

Kim, J.-K., S. E. Bae, S. J. Lee, and M. G. Yoon. 2017. New insight into hybridization and unidirectional introgression between *Ammodytes japonicus* and *Ammodytes heian* (Trachiniformes, Ammodytidae). *PLoS One* 12: e0178001.

Klingenberg, C. P. 2011. MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources* 11: 353–357.

Knudsen, J. T., R. Eriksson, J. Gershenzon, and B. Ståhl. 2006. Diversity and distribution of floral scent. *Botanical Review* 72: 1–120.

Kretzschmar, H., W. Eccarius, and H. Dietrich. 2007. The orchid genera *Anacamptis*, *Orchis*, *Neotinea*. Echinomedia Verlag, Bürgel, Germany.

Lepais, O., R. J. Petit, E. Guichoux, J. E. Lavabre, F. Alberto, A. Kremer, and S. Gerber. 2009. Species relative abundance and direction of introgression in oaks. *Molecular Ecology* 18: 2228–2242.

Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* 20: 229–237.

Martinsen, G. D., T. G. Whitham, R. J. Turek, and P. Keim. 2001. Hybrid populations selectively filter gene introgression between species. *Evolution* 55: 1325–1335.

Meirmans, P. G. 2015. Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* 24: 3223–3231.

Michalski, S. G., and W. Durka. 2015. Separation in flowering time contributes to the maintenance of sympatric cryptic plant lineages. *Ecology and Evolution* 5: 2171–2184.

Milne, R. I., S. Terzioglu, and R. J. Abbott. 2003. A hybrid zone dominated by fertile F1s: maintenance of species barriers in *Rhododendron*. *Molecular Ecology* 12: 2719–2729.

Mitteroecker, P., and P. Gunz. 2009. Advances in geometric morphometrics. *Evolutionary Biology* 36: 235–247.

Mitteroecker, P., P. Gunz, S. Windhager, and K. Schaefer. 2013. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix* 24: 59–66.

Montalvo, A. M., and N. C. Ellstrand. 2001. Nonlocal transplantation and outbreeding depression in the subshrub *Lotus scoparius* (Fabaceae). *American Journal of Botany* 88: 258–269.

Mota, M. R., F. Pinheiro, B. S. S. Leal, T. Wendt, and C. Palma-Silva. 2019. The role of hybridization and introgression in maintaining species integrity and cohesion in naturally isolated inselberg bromeliad populations. *Plant Biology* 21: 122–132.

Neiland, M. R. M., and C. C. Wilcock. 1995. Maximisation of reproductive success by European Orchidaceae under conditions of infrequent pollination. *Protoplasma* 187: 39–48.

Nettel, A., R. S. Dodd, Z. Afzal-Rafii, and C. Tovilla-Hernández. 2008. Genetic diversity enhanced by ancient introgression and secondary contact in East Pacific black mangroves. *Molecular Ecology* 17: 2680–2690.

Paun, O., F. Forest, M. F. Fay, and M. W. Chase. 2009. Hybrid speciation in angiosperms: parental divergence drives ploidy. *New Phytologist* 182: 507–518.

Peakall, R., and P. Smouse. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Peakall, R., and P. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28: 2537–2539.

Pellegrino, G., F. Bellusci, and A. Musacchio. 2009. Genetic integrity of sympatric hybridising plant species: the case of *Orchis italica* and *O. anthropophora*. *Plant Biology* 11: 434–441.

Pellegrino, G., S. D'Emerico, A. Musacchio, A. Scrugli, and S. Cozzolino. 2005. Confirmation of hybridization among sympatric insular populations of *Orchis mascula* and *O. provincialis*. *Plant Systematics and Evolution* 251: 131–142.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Website: https://www.r-project.org/

Raudnitschka, D., C. Henson, and C. Oberprieler. 2007. Introgressive hybridization of *Senecio hercynicus* and *S. ovatus* (Compositae, Senecioneae) along an altitudinal gradient in Harz National Park (Germany). *Systematics and Biodiversity* 5: 333–344.

Rohlf, F. J. 2010a. tpsUtil, version 1.46. Department of Ecology and Evolution, State University of New York at Stony Brook. Website: http://life.bio.sunysb.edu/morph/index.html

Rohlf, F. J. 2010b. tpsDig, version 2.16. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY, USA. Website: http://life.bio.sunysb.edu/morph/index.html

Rohlf, F. J., and D. Slice. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* 39: 40–59.

Rose, F. 1948. *Orchis purpurea* Huds. *Journal of Ecology* 36: 366.

Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener, and S. A. Krawetz [eds.], Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Salzmann, C. C., S. Cozzolino, and F. P. Schiestl. 2007. Floral scent in food-deceptive orchids: species specificity and sources of variability. *Plant Biology* 9: 720–729.

Savriama, Y. 2018. A step-by-step guide for geometric morphometrics of floral symmetry. *Frontiers in Plant Science* 9: 1433.

Schatz, B. 2006. Fine scale distribution of pollinator explains the occurrence of the natural orchid hybrid ×*Orchis bergonii*. *Écoscience* 13: 111–113.

Schatz, B., D. Genoud, J. Claessens, and J. Kleynen. 2020. Orchid–pollinator network in Euro-Mediterranean region: what we know, what we think we know, and what remains to be done. *Acta Oecologica* 107: 103605.

Schatz, B., A. Geoffroy, B. Dainat, J.-M. Bessière, B. Buatois, M. Hossaert-McKey, and M.-A. Selosse. 2010. A case study of modified interactions with symbionts in a hybrid Mediterranean orchid. *American Journal of Botany* 97: 1278–1288.

Scopece, G., A. Musacchio, A. Widmer, and S. Cozzolino. 2007. Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution* 61: 2623–2642.

Scopece, G., S. Cozzolino, and R. M. Bateman. 2010. Just what is a genus? Comparing levels of postzygotic isolation to test alternative taxonomic hypotheses in Orchidaceae subtribe Orchidinae. *Taxon* 59: 1754–1764.

Scopece, G., A. Widmer, and S. Cozzolino. 2008. Evolution of postzygotic reproductive isolation in a guild of deceptive orchids. *American Naturalist* 171: 315–326.

Šmíd, J., J. Douda, K. Krak, and B. Mandák. 2020. Analyses of hybrid viability across a hybrid zone between two *Alnus* species using microsatellites and cpDNA markers. *Genes* 11: 770.

Van der Cingel, N. 1995. An atlas of orchid pollination - European orchids. Balkema, Rotterdam, Netherlands.

Van Hengstum, T., S. Lachmuth, J. J. B. Oostermeijer, H. J. C. M. den Nijs, P. G. Meirmans, and P. H. van Tienderen. 2012. Human-induced hybridization among congeneric endemic plants on Tenerife, Canary Islands. *Plant Systematics and Evolution* 298: 1119–1131.

Verity, R., and R. A. Nichols. 2016. Estimating the number of subpopulations (*K*) in structured populations. *Genetics* 203: 1827–1839.

Viruel, J., A. Haguenauer, M. Juin, F. Mirleau, D. Bouteiller, M. Boudagher-Kharrat, L. Ouahmane, et al. 2018. Advances in genotyping microsatellite markers through sequencing and consequences of scoring methods for *Ceratonia siliqua* (Leguminosae). *Applications in Plant Sciences* 6: e01201.

Viscosi, V. 2015. Geometric morphometrics and leaf phenotypic plasticity: assessing fluctuating asymmetry and allometry in European white oaks (*Quercus*). *Botanical Journal of the Linnean Society* 179: 335–348.

Watano, Y., A. Kania, and N. Tani. 2004. Genetic structure of hybrid zones between *Pinus pumila–P. parviflora* var. *pentaphylla* (Pinaceae) revealed by molecular hybrid index analysis. *American Journal of Botany* 91: 65–72.

Wickham, H. 2009. ggplot2: elegant graphics for data analysis. Springer, NY, NY, USA.

Willems, J. H., and J. Ellers. 1996. Plant performance and population characteristics of *Orchis simia* (Orchidaceae) in two extremes of its distribution area. *Flora* 191: 41–48.

Williams, J. H., W. J. Boecklen, and D. J. Howard. 2001. Reproductive processes in two oak (*Quercus*) contact zones with different levels of hybridization. *Heredity* 87: 680–690.

Yan, L. J., K. S. Burgess, R. I. Milne, C. N. Fu, D. Z. Li, and L. M. Gao. 2017. Asymmetrical natural hybridization varies among hybrid swarms between two diploid *Rhododendron* species. *Annals of Botany* 120: 51–61.

Yan, L. J., K. S. Burgess, W. Zheng, Z. B. Tao, D. Z. Li, and L. M. Gao. 2019. Incomplete reproductive isolation between *Rhododendron* taxa enables hybrid formation and persistence. *Journal of Integrative Plant Biology* 61: 433–448.

Zanella, C. M., C. Palma-Silva, M. Goetze, and F. Bered. 2016. Hybridization between two sister species of Bromeliaceae: *Vriesea carinata* and *V. incurvata*. *Botanical Journal of the Linnean Society* 181: 491–504.

Zeng, Y. F., W. J. Liao, R. J. Petit, and D. Y. Zhang. 2011. Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Molecular Ecology* 20: 4995–5011.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information tab for this article.

Appendix S1. Supplementary methods.
Appendix S2. Supplementary results, including Table S1–5 and Figs. S1–5.
Table S1. Plant DNA samples used to construct genome skim libraries.
Table S2. Samples used to test microsatellite polymorphism.
Table S3. Characteristics and PCR amplification conditions of loci isolated from *Orchis* species.
Table S4. Mean posterior probabilities calculated in NewHybrids for two generations of simulated hybrids and real allopatric populations.
Table S5. Accuracy of discriminant analysis of labellum shape.
Figure S1. Principal component analysis (PC1–PC3) of four allopatric *Orchis* populations, each one representing a different species genotyped with nuclear microsatellites ($n = 30$ in each case).
Figure S2. Thermodynamic integration using RMAVERICK using individuals from allopatric *Orchis* populations and hybrid zones ($N = 695$).
Figure S3. Hybrid indices calculated for individual HYB and ANG populations.
Figure S4. Consensus configurations for labella in *Orchis* populations.
Figure S5. Principal component analysis of labellum shape.

How to cite this article: Bersweden L., J. Viruel, B. Schatz, J. Harland, R. Gargiulo, R. S. Cowan, J. Calevo, A. Juan, J. J. Clarkson, A. R. Leitch, and M. F. Fay. 2021. Microsatellites and petal morphology reveal new patterns of admixture in *Orchis* hybrid zones. *American Journal of Botany*. 108(8): 1388–1404. https://doi.org/10.1002/ajb2.1710